DOCUMENT RESUME

ED 214 161 CS 006 594

AUTHOR Polin, Ruth M.
TITLE A Study of Preceptor Training of Classroom Teachers in Reading Diagnosis.
INSTITUTION Michigan State Univ., East Lansing. Inst. for Research on Teaching.
SPONS AGENCY National Inst. of Education (ED), Washington, D.C.
REPORT NO IRT-RS-110
PUB DATE Nov 81
CONTRACT 400-81-0014
NOTE 48p.; Several pages may be marginally legible.
AVAILABLE FROM The Institute for Research on Teaching, College of Education, Michigan State University, 252 Erickson Hall, East Lansing, MI 48824 ($3.50).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Elementary School Teachers; Inservice Teacher Education; Learning Problems; *Reading Diagnosis; Reading Difficulties; *Reading Research; *Skill Development; *Trainers; *Training Methods

ABSTRACT

A study investigated the effects of differing types of small group instruction on the diagnostic performance of classroom teachers who were inexperienced in reading diagnosis. Fifteen elementary school teachers were divided into three groups and trained for 40 hours by three different preceptors using differing approaches to the same model of reading. Each group gained experience in dealing with either real cases, simulated cases, or both, with instructor feedback, and each was instructed in the use of decision aids to guide interaction in the simulated cases. Student progress was monitored by means of a pretest, midtest, and posttest on a simulated case, and an additional posttest on a case not previously diagnosed. The students diagnosed the cases using a structured form and transferred their diagnoses to a checklist. Statistics were then run on agreement between each student and his or her preceptor, agreement among the students, and agreement among the preceptors. Results showed that student agreement with preceptor and with other students increased in all aspects of a diagnosis; however, transfer of skills to a new case was questionable. Appendixes contain copies of the subject instructions for training sessions, a diagnostic checklist, and an explanation of the statistical analysis used in the study. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED214161

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Research Series No. 110

A STUDY OF PRECEPTOR TRAINING OF CLASSROOM TEACHERS IN READING DIAGNOSIS

Ruth M. Polin

Published By

The Institute for Research on Teaching
252 Erickson Hall
Michigan State University
East Lansing, Michigan 48824

November 1981

CS006594

This work is sponsored in part by the Institute for Research on Teaching, College of Education, Michigan State University. The Institute for Research on Teaching is funded primarily by the Program for Teaching and Instruction of the National Institute of Education, United States Department of Education. The opinions expressed in this publication do not necessarily reflect the position, policy, or endorsement of the National Institute of Education. (Contract No. 400-81-0014)

Institute for Research on Teaching

Teachers' thoughts and decisions are the focus of studies currently under way at Michigan State University's Institute for Research on Teaching (IRT). The IRT was founded in April 1976 with a $3.6 million grant from the National Institute of Education. A new grant obtained in 1981 from the NIE extends the IRT's work through 1984. Funding is also received from other agencies and foundations. The Institute has major projects investigating teacher decision-making, including studies of reading diagnosis and remediation, classroom management strategies, instruction in the areas of language arts, reading, and mathematics, teacher education, teacher planning, effects of external pressures on teachers' decisions, socio-cultural factors, and teachers' perceptions of student affect. Researchers from many different disciplines cooperate in IRT research. In addition, public school teachers work at IRT as half-time collaborators in research, helping to design and plan studies, collect data, and analyze results. The Institute publishes research reports, conference proceedings, occasional papers, and a free quarterly newsletter for practitioners. For more information or to be placed on the IRT mailing list please write to:: The IRT Editor, 252 Erickson, MSU, East Lansing, Michigan 48824.

Co-Directors: Jere E. Brophy and Andrew C. Porter

Associate Directors: Judith E. Lanier and Lee S. Shulman

Editorial Staff:

Janet Flegg, IRT editor
Pat Nischan, assistant editor

A STUDY OF PRECEPTOR TRAINING OF CLASSROOM TEACHERS IN READING DIAGNOSIS[1]

Ruth M. Polin[2]

Clinician agreement on diagnosis, presumed necessary for effective remediation of children's reading difficulties, has been the focus of in-depth investigation by the IRT's Clinical Studies Project for a number of years. The theoretical base for the project is a theory of clinical problem-solving behavior by Elstein and Shulman (Elstein, Shulman, & Sprafka, 1978), which was reformulated for education by Vinsonhaler and Wagner (Vinsonhaler, Wagner, & Elstein, 1977). This theory postulates that problem-solving behavior is a function of clinicians' memories and strategies, which influence the decision-making process as clinicians diagnose cases. Memories and strategies are determined largely by the training clinicians or teachers receive. Clinical decision making in reading has been investigated with several classifications of professionals (reading clinicians, classroom teachers, learning disabilities specialists, and school psychologists) in a series of observational studies from 1977 to 1980, and the results of these studies have been presented by Weinshank (Note 1). There was low diagnostic agreement among all professional groups except school

---

[1]A comprehensive paper on all the training studies will be available at a later date.

[2]Ruth Polin is data-processing coordinator for the Clinical Studies Project.

The author acknowledges the contribution of the following members of the Clinical Studies Project and consultants who assisted in the planning, development of computer systems for data analysis, and operation of the study: John F. Vinsonhaler, Christian Wagner, Annette Weinshank, George Sherman, Gloria Blatt, Doron Gil, Deborah Cureton, Linda Vavrus, and Gwyneth Stewart.

psychologists, whose training and experience differs markedly from that of the others mentioned.

To explore the Instructional Corollary to this Inquiry Theory, which states that clinical performances (e.g., reading diagnosis agreement) can be improved by improvements in clinical memory and strategy, a series of training (application) studies were undertaken. Simply stated, what a clinician retains from his/her educational background about how a child learns to read and what procedures should be followed in assessing the child's mastery of the fundamental elements of reading constitute the clinician's memory and strategy. More intensive, organized training and clinical experience should thus improve the diagnostic decision making of these clinicians and, ultimately, the remediation of a child's reading disability.

The first of these application studies was conducted in 1977 with the students of the graduate reading diagnosis course at Michigan State University (Sherman, Weinshank, & Brown, Note 2). In 1979, another study was conducted in the same setting, but with the additional use of a model of reading for instruction and of decision aids (diagnostic write-up forms and checklists). Simulated cases developed by the Clinical Studies Project were used in all studies. Simulated cases are simply collections of information about children with reading problems. They are based on real children in grades three through nine who attended the Michigan State University Reading Clinic and whose reading problems were considered to be representative of those frequently noted in public schools. A detailed description of these cases is presented by Weinshank (Note 1). The 1979 study showed the efficacy of using diagnostic decision aids in channeling a clinician's thinking along specified, orderly lines; this resulted in superior inter-clinician

agreement (Gil, Polin, Vinsonhaler, & VanRoekel, Note 3).

That result prompted further investigation of the effect of training on diagnostic agreement. In 1980, the study reported here was conducted. In this study, I and my colleagues on the Clinical Studies Project investigated the effect of differing types of small group instruction on the diagnostic performance of classroom teachers who were inexperienced in reading diagnosis.

Objectives

The objectives of this application study were to attempt to answer the following questions.

1. Since the 1979 application study showed the positive effect of systematized training in reading diagnosis, would different content and training methods also have a positive effect?

2. Does practice with simulated cases have a greater effect on diagnostic performance than the necessarily limited practice with real children?

3. If limited training results in higher diagnostic agreement, can even higher agreement be expected with more extensive training?

Methods

Subjects

Ten classroom teachers with little or no training or experience in reading diagnosis, drawn from elementary schools of the Lansing, Michigan school district, and five students enrolled in the Michigan State University graduate-level course in reading diagnosis, also with minimal experience in reading diagnosis, participated in the study. The students were divided into three preceptor training groups, each with a different instructor (preceptor). Each instructor had a somewhat different approach to teaching reading diagnosis. However, all groups

were instructed for a minimum of 30 hours plus 10 hours of extra practice time with (1) the Model of Reading and Learning to Read (MORAL), a systematized cause and effect approach to reading difficulty diagnosis (Sherman, Note 4); (2) either simulated cases, real cases, or both, with instructor feedback; and (3) decision aids, which guided the interaction of those using the simulated cases. Progress was monitored by means of pre-, mid-, and posttests on a simulated case, and an additional posttest on a case not previously diagnosed. Five simulated cases were used; one subject from each preceptor training group was tested on each case.

Instructional Design

The vital signs of reading are the basis for group instruction in diagnosis of cases of reading difficulty. These vital signs indicate the reader's "health," much like medical vital signs indicate a patient's health. They are as follows:

1. *Instant Word Recognition*--the ability to recognize words without hesitation. It is measured by size of sight word vocabulary, relative to a child's grade placement in school.

2. *Decoded Word Recognition*--the ability to recognize unfamiliar words through use of graphemic similarities.

3. *Meaning Vocabulary*--the scope of words that denote meaningful relationships to the reader.

4. *Oral Reading*--the ability to read aloud with fluency and inflection.

5. *Reading Comprehension*--the ability to understand and put into meaningful perspective material that is read.

6. *Listening Comprehension*--the ability to understand and put into meaningful perspective material that is heard.

7. *Attention/Motivation*--the ability to activate and maintain concentration on the task at hand.

The preceptor for Group 1 used four vital signs. He merged Vital Signs 3 and 4, calling them Fluent Text Segmentation, and 5 and

6, calling them Retentive Comprehension. He considered Vital Sign 7 a learning effecting factor, rather than a vital sign. (A learning effecting factor is something that directly influences learning.) Group 1 stressed diagnosis according to performance level on the four vital signs; the other two groups used all seven as separate vital signs.

Classroom Instruction

For Groups 2 and 3, the formal classroom instruction in reading diagnosis was conducted in three-hour blocks weekly with an additional 1½ hours per week spent in diagnosing computer-based simulated cases (as opposed to the manually-based ones used for the test sessions). Computer-based simulated cases contained the same information as the manual simulated cases but were used in conjunction with a computer terminal, which had test information stored on disks. This information could be readily retrieved for instant display without the use of the cumbersome file box (manual). Any material that could not be stored on the disks (e.g., audio recordings, actual test booklets) were contained in a loose-leaf study guide. After interacting with a simulated case, students filled out the decision-aid diagnosis sheets. Each student then translated his/her diagnosis to a standardized checklist, indicating whether the case showed adequacies or inadequacies in the seven vital signs and their effecting factors as postulated by the Model of Reading and Learning to Read. (This procedure is described in detail in the next section of this paper.) Students in Group 1, who used real cases rather than the simulated ones, did not use the checklist for their cases. Instead, the preceptor analyzed in class the real cases diagnosed by each student, and this in-class analysis provided practice comparable to that of

the other two groups. Some of the instructional differences across the three training groups are summarized in Table 1.

The diagnosis and remediation chart used in Groups 2 and 3 is shown, for one vital sign, in Figure 1. The same format was used for each of the seven vital signs. Group 1's flowchart also outlined steps to be followed in diagnosing cases of reading difficulty according to vital signs and effecting factors.

Test Sessions

At each test session, subjects were presented with a manually-based simulated case. The students received written and oral instructions on how to find the information and how to complete the test session (see Appendix A).

After receiving instructions, subjects observed the initial contact information about the case, which included a short summary about the child's reading performance. The subjects were then given 45 minutes to collect as many cues (items of information) about the case as they wished. These cues were later recorded as part of their data for the test.

At the end of 45 minutes, each student was asked to write a diagnosis of the case, using the decision aid, a sample of which is shown in Figure 2, based on the child's adequacies or inadequacies on the seven vital signs. The students were given 30 minutes to do this.

After completing their diagnostic write-ups, students were required to match their written diagnoses with the diagnostic categories listed on a checklist. This transfer, made after the students wrote their diagnoses, was used to establish a standardized vocabulary for analysis. The students did not see the checklist while they were writing their diagnoses.

Table 1

Differences and Similarities Across Preceptor Groups

Features	Group 1	Group 2	Group 3
Out of class practice on simulated cases	no	yes	yes
Practice with real children	yes	yes	yes
Use of Informal Reading Inventory in diagnosis	no	yes	no
Use of weekly logs for feedback	yes	no	yes
Number of vital signs studied	4	7	7
Decision aids used	flowchart[a]	diagnosis and remediation charts[b]	diagnosis and remediation charts[b]
Text used (for outside reading)	Ekwall[c]	Pearson & Johnson[d] MORAL for preceptors[e]	Pearson & Johnson[d] MORAL for preceptors[e]

[a]Developed by George Sherman.

[b]Developed by Deborah Cureton, Linda Patriarca, and Gwyneth Stewart.

[c]Ekwall, E. Diagnosis and remediation of the disabled reader. Boston, Mass.: Allyn & Bacon, 1976.

[d]Two texts:
Pearson, P.D., & Johnson, D.D. Teaching reading comprehension. New York: Holt, Rinehart & Winston, 1978.

Johnson, D.D., & Pearson, P.D. Teaching reading vocabulary. New York: Holt, Rinehart & Winston, 1978.

[e]Vinsonhaler, Weinshank, Cureton, & Blatt (Note 5).

Instant Word Recognition

How to measure the process?	What is the process being measured?	Factors That Effect the Process	Possible Problems
(performance on)	(state of)		
Sight Word List	Instant Word Recognition		
Oral Reading - error analysis	def. the ability to see and say a word accurately if asked to do so quickly; often referred to as sight vocabulary sight words	A. Visual Discrimination of Words	A. Confusion of look-alike words
		B. Visual Memory of Words	B. Inability to remember words encountered in print
		C. Semantic representations (meanings) of words	C. Word meanings Lacking
		D. Association of graphemic with phonologic representation of words	D. Inability to pronounce a written word
		E. Association of graphemic with semantic representation of words	E. Inability to attach meaning to a written word

---

Learning Process	Learning Process Effecting Factors	Possible Problems
Learning: change in one's semantic net induced by effective practice	Amount and conditions of effective practice	Defective practice
	A. Attention of the Learner	A. Learner's attention unfocused
	1. motivation to read	1. Lack of motivation
	a. internal	a. no interest to read
	b. external	b. environment not conducive to reading
	2. reader's self-concept	2. poor self-concept
	3. match between interest abilities and materials	3. poor match between interest materials and readiness skills
	B. Relevance (transferability) of the practice task to the learning task	B. Irrelevant task
	C. Learner's correct preception of the task	C. Incorrect perception of the task
	1. visual acuity	1. poor vision
	2. auditory acuity	2. poor hearing
	3. clarity of the learning objective	3. objectives unclear
	4. clarity of the criteria for success	4. criteria unclear
	D. Corrective feedback	D. Defective feedback
	1. attention feedback loop	1. negative feedback loop
	2. dependence of feedback on performance	2. feedback inappropriate for the task

Figure 1. Diagnostic/Remedial chart for first vital sign of reading.

Application Study 1980 Diagnostic / Remedial Record

Case Name _____
Your Name _____
Date _____

Does the student have a problem with INSTANT WORD RECOGNITION?

(Circle One) Yes No

On what basis was this decision made?

If no, then continue with the next problem area on page 3.

If yes, describe the important factors that have contributed to this problem. For each factor, suggest remedial procedures for its improvement. Continue on the next page if required.

1.Describe one factor contributing to the problem with Instant Word Recognition.

Suggest remedial procedures for alleviating this factor.

2.Describe another factor contributing to the problem with Irstant Word Recognition.

Suggest remedial procedures for alleviating this factor.

3.Describe another factor contributing to the problem with Instant Word Recognition.

Suggest remedial procedures for alleviating this factor.

Figure 2. Decision aid used by students in diagnosing cases.

The checklist listed the seven vital signs as major categories, requiring a decision as to adequate or inadequate performance for each. Subsets under each vital sign included specifics important to the sign and decisions as to whether those performances were adequate or inadequate. Under each vital sign, an "other" category was listed to accommodate those diagnostic statements on the student's diagnoses that could not be translated into existing categories. In addition, some learning effecting factors were listed separately at the end of the checklist. (The complete checklist is contained in Appendix B.)

At the end of the session. the written diagnosis, record form for cues, and checklist of each student were collected and filed. Each student was assigned a subject number using a random-order table selection, and all data were processed.

The first test session was held prior to any group meetings (pretest). Identical procedures were followed for the midtest (approximately five weeks later) and the posttest (at the end of the 10-week session). In all these tests, students diagnosed the same case, thus enabling a progress profile to be established. A week after the posttest (Posttest 1), another posttest was given (Posttest 2). In this test, students diagnosed a different simulated case, one which they had never seen, and, in addition, were allowed to use the written diagnosis and remediation charts and flowcharts, which were part of their in-class training and practice on cases. In the previous test sessions, the charts were not available; thus, the students had to internalize the material on the charts as a guide to writing their diagnoses.

Results

Phi correlation and Porter statistic were calculated to find the inter-student agreement, the student-preceptor agreement, and the inter-preceptor agreement. The latter indicated the agreement of experts on each case and served as a standard against which student performance could be measured. (For a discussion of these statistics, see Appendix C.)

A member of the Clinical Studies Project checked that all transfers from the written diagnosis to the checklist were accurate, being careful not to tamper with intent. Such errors as marking an item as both adequate and inadequate and omitting vital signs indicated as problem areas on the written diagnosis were found in both student and preceptor checklists and were corrected. All analysis was then made on these amended diagnoses.

Total Diagnosis

The results of the total diagnosis, which includes both vital signs and effecting factors, are presented in Table 2. The table contains the agreement of each student with his/her preceptor for all three groups on all five cases. There were 128 categories of statements--64 each of adequacies and inadequacies.

Vital Signs and Effecting Factors

Because every diagnosis consists of two components, vital signs

Table 2

Agreement of Student With His/Her Preceptor on Complete, Corrected Diagnosis

Key
PTG = Preceptor Training Group
Ø = Phi Coefficient
P = Porter Statistic

Case	Pretest						Midtest						Posttest 1						Posttest 2					
	PTG 1		PTG 2		PTG 3		PTG 1		PTG 2		PTG 3		PTG 1		PTG 2		PTG 3		PTG 1		PTG 2		PTG 3	
	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P
1	.62	.46	.23	.19	.38	.29	.64	.46	.49	.41	.46	.36	.49	.34	.58	.50	.41	.32	.14	.14	.17	.20	.29	.24
2	.30	.26	.38	.33	.26	.21	.34	.29	.50	.44	.39	.31	.39	.32	.59	.55	.43	.32	.30	.26	.56	.55	.29	.24
3	.34	.28	.46	.36	.46	.33	.19	.19	.48	.39	.21	.18	.14	.16	.44	.39	.36	.28	.44	.35	.37	.35	.12	.13
4	.50	.35	.38	.26	.41	.32	.71	.59	.33	.29	.50	.40	.50	.40	.36	.31	.53	.43	.31	.26	.44	.39	.32	.26
5	.34	.27	.27	.22	.34	.26	.52	.41	.35	.27	.31	.26	.64	.52	.35	.30	.35	.28	36	.28	.35	.31	.21	.19
Grand Mean	.42	.32	.34	.27	.37	.28	.48	.39	.43	.36	.37	.31	.43	.35	.46	.41	.42	.33	.31	.26	.38	.36	.25	.22
St. Dev.	.14	.08	.09	.07	.08	.05	.21	.15	.08	.08	.12	.09	.19	.13	.12	.11	.07	.06	.11	.08	.14	.13	.08	.05
Mean of Groups	Ø = .38 P = .29						Ø = .43 P = .35						Ø = .44 P = .36						Ø = .31 P = .28					

and effecting factors, the diagnoses were analyzed for each component part. Table 3 contains results on vital signs only (domain = 14), and Table 4 presents effecting-factors results (domain = 112). Table 5 presents inter-student agreement for the three types of diagnoses.

Inter-Preceptor Agreement

Table 6 presents the inter-preceptor agreement under all three conditions (total diagnosis, vital signs only, and effecting factors only). The highest agreement appears to be between the preceptor of Group 1 and the preceptor of Group 3; the lowest agreement was between Preceptors 1 and 2. Preceptors 1 and 3 were trained and practice at Michigan State University. Preceptor 2 trained at both Michigan State University and other universities and is currently on the staff at another university. The results may also show the presence of a ceiling effect, that is, a limit under present analytical treatment to which agreement can be otained.

Cues

In Table 7, the student-preceptor agreement on cues selected is presented. Agreement appears to be lower than that on diagnostic statements. Likewise, agreement among preceptors on which cues to select was also lower than their diagnostic agreement (Table 8) and again, the best agreement was between Preceptors 1 and 3. In addition, these two preceptors agreed more closely on the number of cues to

Table 3

Agreement of Student with his/her Preceptor on Vital Signs

Key

PTG = Preceptor Training Group
Ø = Phi Coefficient
P = Porter Statistic

Case	Pretest						Midtest						Posttest 1						Posttest 2					
	PTG 1		PTG 2		PTG 3		PTG 1		PTG 2		PTG 3		PTG 1		PTG 2		PTG 3		PTG 1		PTG 2		PTG 3	
	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P
1	.71	.75	.71	.75	.43	.56	1.00	1.00	.71	.75	1.00	1.00	.71	.75	.43	.56	.43	.56	.29	.44	.43	.56	.43	.56
2	.43	.56	.71	.75	.71	.75	.43	.56	.43	.56	.71	.75	.14	.40	1.00	1.00	.71	.75	.14	.40	1.00	1.00	.14	.40
3	.71	.75	.71	.75	.14	.40	-.14	.27	.43	.56	.43	.56	-.14	.27	.43	.56	.14	.40	.43	.56	.43	.56	-.14	.27
4	.43	.56	.43	.56	.71	.75	.71	.75	.43	.56	.71	.75	.43	.56	.43	.56	.71	.75	.43	.56	.43	.56	.43	.56
5	.43	.56	.71	.75	-.14	.27	.71	.75	.71	.75	.43	.56	.71	.75	.71	.75	.43	.56	1.00	1.00	.71	.75	.43	.56
Grand Mean	.54	.64	.65	.71	.37	.55	.54	.67	.54	.64	.66	.72	.37	.55	.60	.69	.48	.60	.46	.59	.60	.69	.26	47
Std. Dev.	.15	.10	.13	.08	.37	.21	.43	.27	.15	.10	.24	.18	.37	.21	.25	.19	.24	.15	.33	.24	.25	.19	.26	.13
Mean of Groups	Ø = .52 P = .63						Ø = .58 P = .68						Ø = .48 P = .61						Ø = .44 P = .58					

Table 4

Agreement of Student with his/her Preceptor on Effecting Factors

Key

PTG = Preceptor Training Group
Ø = Phi Coefficient
P = Porter Statistic

Case	Pretest						Midtest						Posttest 1						Posttest 2					
	PTG 1		PTG 2		PTG 3		PTG 1		PTG 2		PTG 3		PTG 1		PTG 2		PTG 3		PTG 1		PTG 2		PTG 3	
	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P
1	.39	.22	.00	.04	.12	.09	.45	.24	.37	.29	.14	.11	.31	.16	.56	.46	.31	.21	.08	.07	.02	.09	.17	.13
2	.19	.16	.22	.19	.01	.05	.29	.22	.48	.40	.19	.15	.40	.27	.48	.44	.39	.26	.18	.15	.51	.46	.20	.16
3	.16	.14	.27	.19	.27	.14	.18	.14	.42	.32	.08	.00	.16	.13	.34	.29	.22	.15	.25	.19	.30	.26	.03	.05
4	-.03	.00	.29	.14	.17	.13	.70	.53	.29	.23	.38	.26	.40	.29	.32	.26	.30	.21	.20	.16	.24	.17	.27	.17
5	.16	.13	.24	.18	.28	.18	.17	.13	.42	.29	.14	.12	.43	.31	.40	.32	.17	.14	.24	.17	.26	.24	.05	.08
Grand Mean	.17	.13	.20	.15	.17	.12	.36	.25	.40	.29	.15	.13	.34	.23	.42	.35	.28	.19	.19	.15	.27	.24	.14	.12
Std. Dev.	.15	.08	.12	.06	.11	.05	.22	.16	.07	.07	.16	.09	.11	.08	.10	.09	.09	.05	.07	.05	.17	.14	.10	.05
Mean of Groups			Ø = .18 P = .13						Ø = .30 P = .22						Ø = .35 P = .26						Ø = .20 P = .17			

Table 5

Inter-Student Correlation

Key

Ø = Phi Correlation
P = Porter Statistic

Total Diagnosis

Simulated Case	Pretest Ø	Pretest P	Mid test Ø	Mid test P	Post test 1 Ø	Post test 1 P	Post test 2 Ø	Post test 2 P
1	.26	.21	.30	.26	.37	.33	.41	.37
2	.29	.25	.34	.31	.45	.45	.41	.34
3	.37	.25	.31	.27	.40	.34	.38	.31
4	.46	.35	.50	.42	.43	.35	.35	.32
5	.31	.25	.35	.29	.47	.39	.48	.42
$\bar{X}$	.34	.26	.36	.31	.42	.37	.41	.35
S.D.	.08	.05	.08	.06	.04	.05	.05	.04

Vital Signs Only

Simulated Case	Pre test Ø	Pre test P	Mid test Ø	Mid test P	Post test 1 Ø	Post test 1 P	Post test 2 Ø	Post test 2 P
1	.24	.46	.43	.57	.62	.69	.72	.74
2	.81	.83	.43	.57	1.00	1.00	.62	.71
3	.24	.46	.42	.56	.62	.69	.62	.70
4	.43	.57	1.00	1.00	.43	.57	.43	.57
5	.24	.45	.62	.69	.62	.69	.81	.83
$\bar{X}$	.39	.55	.58	.68	.66	.73	.64	.71
S.D.	.24	.16	.25	.19	.21	.16	.14	.09

Effecting Factors

Simulated Case	Pre test Ø	Pre test P	Mid test Ø	Mid test P	Post test 1 Ø	Post test 1 P	Post test 2 Ø	Post test 2 P
1	-.04	.02	.21	.18	.26	.23	.32	.29
2	.11	.11	.27	.24	.42	.39	.31	.24
3	.13	.07	.25	.20	.22	.20	.29	.21
4	.14	.10	.36	.28	.37	.27	.27	.24
5	.08	.08	.18	.16	.26	.21	.32	.28
$\bar{X}$	.08	.08	.25	.21	.31	.26	.30	.25
S.D.	.07	.04	.07	.05	.08	.08	.02	.03

Table 6

Preceptor/Preceptor Agreement on Diagnosis

Ø = Phi Coefficient

P = Porter Statistic

P1 = Group 1 Preceptor
P2 = Group 2 Preceptor
P3 = Group 3 Preceptor

	Total Dx						Vital Signs						Effecting Factors					
	P1 vs. P2		P1 vs. P3		P2 vs. P3		P1 vs. P2		P1 vs. P3		P2 vs. P3		P1 vs. P2		P1 vs. P3		P2 vs. P3	
Case	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P
1	.29	.23	.68	.56	.37	.29	.43	.56	.71	.75	.71	.75	.14	.10	.48	.33	.19	.14
2	.26	.24	.51	.40	.42	.32	.14	.40	.43	.56	.71	.75	.28	.21	.45	.33	.34	.22
3	.55	.43	.82	.72	.51	.38	.43	.56	1.00	1.00	.43	.56	.52	.37	.51	.36	.45	.28
4	.33	.28	.57	.46	.57	.45	.14	.40	.43	.56	.71	.75	.29	.21	.56	.42	.41	.27
5	.34	.24	.38	.29	.32	.26	1.00	1.00	.43	.56	.43	.56	.23	.16	.06	.07	.23	.16
Grand Mean	.35	.28	.59	.49	.44	.34	.43	.58	.60	.69	.60	.67	.29	.21	.41	.30	.32	.21
Std. Dev.	.11	.08	.17	.16	.10	.08	.35	.25	.25	.19	.15	.10	.14	.10	.20	.13	.11	.06
Mean Across Preceptors	Ø = .46 P = .37						Ø = .54 P = .64						Ø = .34 P = .24					

Table 7

Agreement of Student With His/Her Preceptor on Cues Collected

Key

PTG = Preceptor Training Group
Ø = Phi Coefficient
P = Porter Statistic

Case	Pretest						Midtest						Posttest 1						Posttest 2					
	PTG 1		PTG 2		PTG 3		PTG 1		PTG 2		PTG 3		PTG 1		PTG 2		PTG 3		PTG 1		PTG 2		PTG 3	
	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P	Ø	P
1	.00	.09	.30	.35	.08	.13	.27	.21	.08	.20	.08	.13	.27	.22	.28	.37	.13	.14	.37	.27	.20	.35	.14	.16
2	.02	.10	.19	.28	-.02	.09	.19	.21	.04	.14	.18	.18	.20	.21	.03	.19	.13	.15	.13	.17	.11	.28	-.11	.05
3	.11	.14	.11	.21	-.16	.02	.70	.58	.09	.21	.20	.17	.45	.32	.01	.20	.02	.06	.52	.41	.04	.17	-.02	.05
4	-.18	0	.06	.17	-.02	.07	.28	.24	.12	.31	.10	.12	.02	.10	.15	.28	.20	.17	.14	.17	.12	.29	.04	.10
5	.09	.18	-.14	.12	.13	.13	.02	.09	-.04	.18	.17	.16	.20	.22	.16	.20	.14	.11	.08	.16	.00	.21	.19	.17
Grand Mean	.01	.10	.10	.23	.00	.09	.29	.27	.06	.21	.15	.15	.23	.21	.13	.25	.12	.13	.25	.24	.09	.26	.05	.11
St. Dev.	.11	.07	.16	.09	.11	.05	.25	.18	.06	.06	.05	.03	.15	.08	.11	.08	.07	.04	.19	.11	.08	.07	.12	.06
Mean of Groups	Ø = .04 P = .14						Ø = .17 P = .21						Ø = .16 P = .20						Ø = .13 P = .20					

Table 8

Inter-Preceptor Agreement on Cues Collected

Case	P1 vs. P2 Ø	P1 vs. P2 P	P1 vs. P3 Ø	P1 vs. P3 P	P2 vs. P3 Ø	P2 vs. P3 P
1	.08	.12	.30	.24	.22	.21
2	.22	.22	.50	.39	.30	.24
3	.02	.09	.28	.21	.07	.10
4	.09	.16	.27	.21	.26	.18
5	.20	.23	.34	.27	.36	.27
Grand Mean	.12	.16	.34	.26	.24	.20
Std. Dev.	.08	.06	.09	.07	.11	.07
Mean Across Preceptors			Ø = .23	P = .21		

Key

Ø = Phi Coefficient
P = Porter Statistic
P1 = Group 1 Preceptor
P2 = Group 2 Preceptor
P3 = Group 3 Preceptor

collect in order to diagnose a case. Preceptor 1 collected 14 cues, Preceptor 2 collected 29 cues, and Preceptor 3 collected only 11.

Changes in Diagnostic Agreement

Figures 3 - 6 show the changes in diagnostic agreement for all students, regardless of group attendance (N=15), as they progressed through the course. The pre- to posttest changes primarily show growth in both student-preceptor and inter-student agreement. However, for the second posttest, when students were faced with a new case, student-preceptor agreement fell off sharply, whereas inter-student agreement remained constant.

Discussion

Differences in performance among students of the three groups could not be related solely to their presence in different training groups (i.e., the differences involved in their training). There were only five students in each group, each student tested on a different simulated case. Differences in simulated-case difficulty and in student ability and background were among the factors that could have influenced performance on diagnosis more than the details of preceptor training. Therefore, conclusions need to be drawn on mean statistics across all three groups (15 subjects) rather than on the limited sample of single preceptor training groups.

Certain statements can be made about the efficacy of training. Results similar to that of the previous study (Gil, Polin, Vinsonhaler, & VanRoekel, Note 3), show that agreement among students and with preceptors was aided by the use of decision aids in diagnoses in all four test sessions. Earlier observational studies (Weinshank, Note 1) involving even experienced reading clinicians in which no such decision

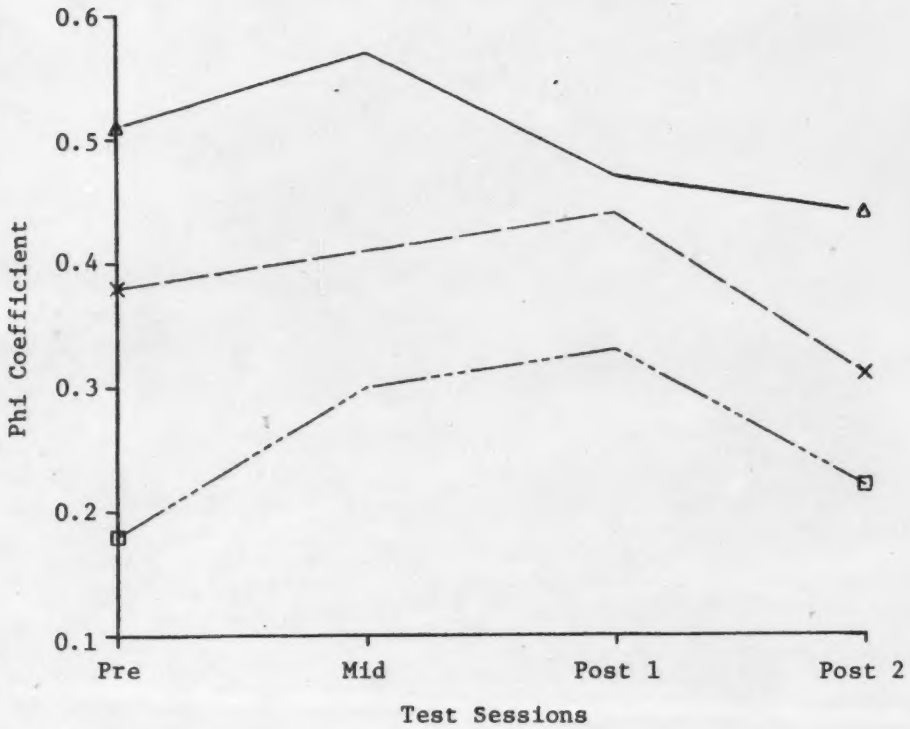


Figure 3. Changes in mean student/preceptor agreement across all four testing sessions (Phi Coefficient).

Legend

△ Vital Signs

X Total Dx

□ Effecting Factors

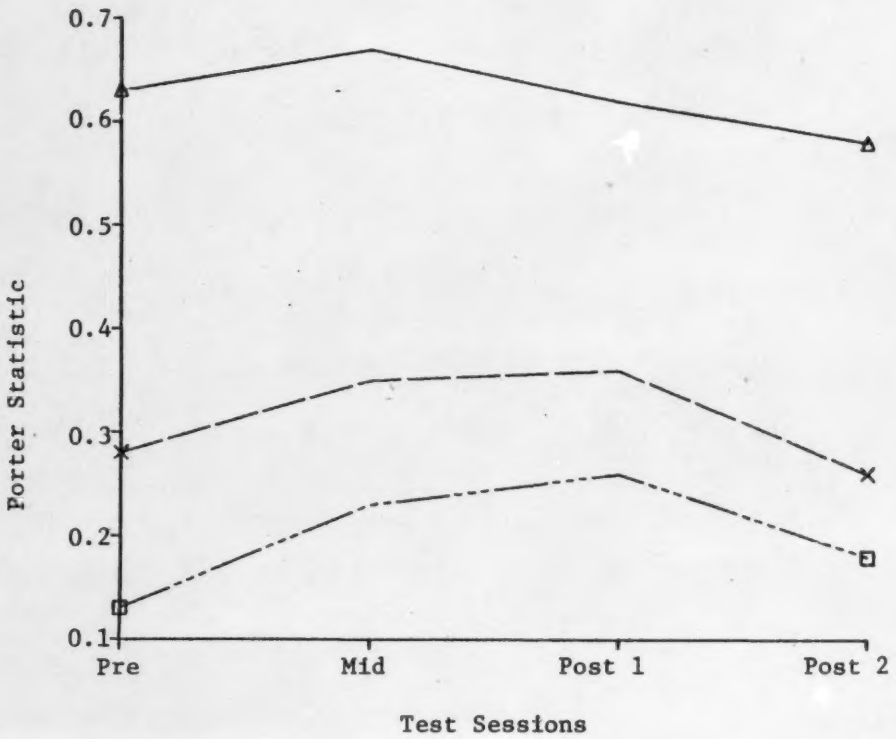


Figure 4. Changes in mean student/preceptor agreement across all 4 testing sessions (Porter Statistic).

Legend

△ Vital Signs _____

X Total Dx _ _ _ _ _

□ Effecting Factors . . . . .

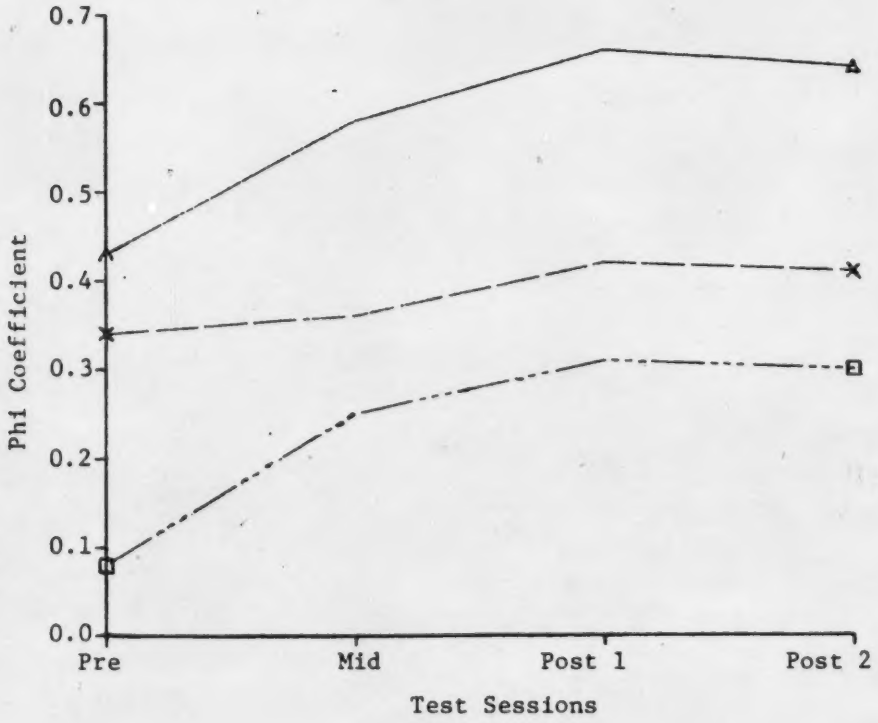


Figure 5. Changes in mean student/student agreement across all four testing sessions (Phi Coefficient).

Legend

△ Vital Signs _____

X Total Dx _ _ _ _ _ _ _ _

□ Effecting Factors _ _ _

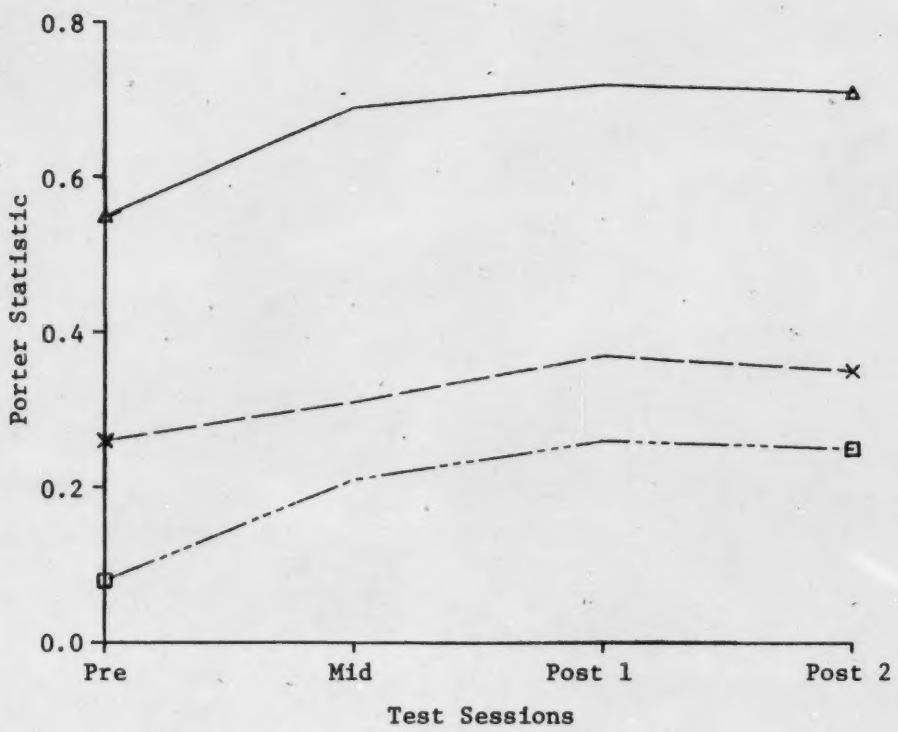


Figure 6. Changes in mean student/student agreement across all four testing sessions (Porter Statistic).

Legend

△ Vital Signs _____

X Total Dx __ __ __ __

□ Effecting Factors _ _ _ _ _

aids were used resulted in considerably lower diagnostic agreement. Two conclusions confirming those of the previous training study are (1) that the use of decision aids appears to be of value in obtaining higher diagnostic agreement, and (2) that student performance does improve with training. It would appear that perhaps additional practice with simulated cases would sharpen students' diagnostic skills. However, the limited inter-preceptor agreement seems to indicate that, at least under the present testing conditions, further increases in diagnostic agreement may not be attainable. For example, the preceptors diagnosed the cases only once and this diagnosis was used for comparison with all three student diagnostic sessions. Thus, there is no data on intra-preceptor agreement--how closely preceptors agree with themselves when diagnosing a case a second or third time.

Student-preceptor agreement was not maintained when the students were given a new case to diagnose (Posttest 2). However, the inter-student agreement remained at the same level. This may mean that somehow the students were using strategies not taught or intended by the preceptors. This poses the question of whether the skills obtained during training can be maintained and generalized. Will classroom teachers continue to use the organization and model in their classrooms? Interviews with students at the end of the study elicited very favorable response to the training and a desire on their part to incorporate their new knowledge into their classroom operations. For example, analysis of the logs kept by the students in Group 3 showed students developing increasing confidence in their ability to diagnose reading difficulties cases. However, if we use

student-preceptor agreement as the criterion, the results of Posttest 2 do not appear to warrant such confidence.

In summary, students do appear to gain confidence as well as precision with training on a model, use of diagnostic aids, and practice with feedback. The latter is most practically accomplished with simulated cases both in terms of time required and control of feedback. The goal of reaching and maintaining a high level of precision (and accuracy if agreement with preceptors is an indicator) poses a need for more than the current 10 - 14 weeks of training now in use by most teacher education institutions, and, more importantly, more exposure to a variety of cases for practice. In addition, it would be valuable to follow students' performance to see if gains made can be maintained and translated into meaningful performance in the classroom.

Reference Notes

1. Weinshank, A.B. Investigations of the diagnostic reliability of reading specialists, learning disabilities specialists, and classroom teachers: Results and implications (Research Series No. 88). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1980.

2. Sherman, G.B., Weinshank, A.B., & Brown, S. Training reading specialists in diagnosis (Research Series No. 31). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1979.

3. Gil, D., Polin, R., Vinsonhaler, J., & Van Roekel, J. The impact of training on diagnostic consistency (Research Series No. 67). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1979.

4. Sherman, G.B. A Model of Reading and Learning to Read. Institute for Research on Teaching, Michigan State University, in progress.

5. Vinsonhaler, J.F., Weinshank, A.B., Cureton, D.B., & Blatt, G. A Model of Reading and Learning to Read. Institute for Research on Teaching, Michigan State University, unpublished paper, 1980.

References

Elstein, A., Shulman, L., & Sprafka, S. Medical problem solving: An analysis of clinical reasoning. Cambridge, Massachusetts: Harvard University Press, 1978.

Vinsonhaler, J.F., Wagner, C.C., & Elstein, A.S. The Inquiry Theory: An information processing approach to clinical problem solving. In Medinfo 1977, proceedings of the Second World Conference in Medical Informatics. Amsterdam, Holland: North Holland Publishing Co., 1977.

Appendix A

Subject Instructions for Observational Sessions

AS80/SPRING 1980 3/15/80 1
Subject Instructions for Observational Sessions

The purpose of the Diagnostic Performance Session is both to instruct you and to evaluate your progress as a diagnostician of reading disabilities. The Diagnostic Performance Session requires you to diagnose a simulated case (SIMCASE) of reading disability. Our SIMCASES are 'simulated' only in the sense that the real child is not present. The data in the SIMCASE was obtained from real children with real reading disabilities.

MAIERIALS The materials you need for this diagnostic performance session are as follows:

(1) The SIMCASE including:

- The List of Available Information
- The Referral Information (Initial contact Folder)
- The Box of Materials

(2) The Diagnostic / Remedial Record
(3) An Audio Tape Cassette Player

If you do not have any of these items signal your instructor.

PROCEDURES
To review, the procedures for the observational session will be as follows: following:

(1) Read the initial contact(referral) information. Look at the sketch of the child and listen to the taped interview;
(2) Diagnose the SIMCASE by collecting whatever further data you wish;
(3) Write up your diagnosis and remediation using the Diagnostic / Remedial Record

This has been a general description of the Diagnostic Performance Session. If you have any questions, signal your instructor.

HOW_IO_USE_IHE_DIAGNOSIIC_FORM Pick up the Diagnostic / Remedial Record and look it over for a few minutes.

The form provides a guide for writing your diagnostic/remedial judgments. As you complete each page, write the case name, your name and the date in the spaces provided. The form is divided into seven parts corresponding to important reading activities. For each part, state your diagnostic decision by circling YES if the reading activity is problematic and NO otherwise. Then, describe the basis on which the decision was made. Finally, if the area is problematic, describe any factors that have contributed to the problem by 1)writing down the factor, and 2) providing remedial suggestions for the factor.

HOW_TO_USE_A_SIMCASE
Let's now consider the specifics of how to use a SIMCASE. The case consists of the following items:

(1) The Initial Contact (Referral) Information
(2) The List of Available Information
(3) The Box of Materials

The Initial_Contact_Information provides some basic introductory information about the child. It includes a referral statement, a sketch of the child, and a taped interview.

The Information_Inventory is the list of materials collected about the case. Each item of information is described by a KEYWORD that refers to its title and location in the box (e.g.,keyword for Durrell Oral Reading test booklet is DUR3).

Now open the SIMCASE_Box.
Note the KEYWORDS on the folders. These keywords correspond to those in the Information_Inventory. To find information in the SIMCASE--locate the proper KEYWORD in the Inventory. Use this KEYWORD to find the proper folder in the box.
As in the case of DUR3, find the large folder entitled DURRELL. Within that folder are smaller colored folders in which will be found the specific item DUR3. Only remove the desired information; not the colored folder in which it is located.

Now that you have practiced, you can begin to diagnose and remediate your SIMCASE.

Begin by opening the Initial Contact folder. Examine the picture and the information, and listen to the tape.

When you have finished with the Initial Contact material, signal your instructor. After he or she responds you may begin collecting information to diagnose your case. You will have 45 minutes in which to collect information. The proctor will tell you when your time is up.

You will have an additional 30 minutes in which to fill out the Diagnostic / Remedial Record and 15 minutes to transfer the diagnosis to a Diagnostic Checklist described later.

Please_do_not_return_any_of_the_materials_to_the_box. Leave_them_on the_table.

If you have questions or problems at any time, signal your instructor. Relax. There are no grades to be given in this Study.

Appendix B

Diagnostic Checklist

Application Study 1980 Diagnostic Checklist 1

Case Name _____
Your Name _____
Date _____

The purpose of the Diagnostic checklist is to translate your diagnosis as written on the Diagnostic/Remedial Record into a standard vocabulary. To that end, please be sure to include only those categories that you have WRITTEN down on the Diagnostic/Remedial Record. Even if you know the student's status on every statement in the checklist, we are only interested in having a translation of the ones you included in your written diagnosis.

Examine your written diagnosis as recorded on the Diagnostic/Remedial Record. For each factor on the Record that describes a FACTOR contributing to a problem, do the following:

1)Circle the factor on the Diagnostic/Remedial Record
2)Number the factor, starting at 1 for the first one and going up from there for each successive factor
3)Locate the statement on this checklist that most accurately characterizes the factor.
4)Place the factor number next to the statement on the checklist.
5)Continue for all factors on your Diagnostic/Remedial Record

An example of this process is given below.

The Diagnostic / Remedial Record is:

1.Describe one factor contributing to the problem with Instant Word Recognition. 1 (poor basic sight words)

Suggest remedial procedures for alleviating this factor.

flash card practice

2.Describe another factor contributing to the problem with Instant Word Recognition. 2 (can't discriminate letters)

Suggest remedial procedures for alleviating this factor.

discrimination through faded cueing training

etc.

And the Diagnostic Checklist is:

1_____Instant Word Recognition Adequate
2_____Instant Word Recognition Inadequate
3_____ Basic Sight Words Adequate
4__1__ Basic Sight Words Inadequate

.

10__2__ Visual Discrimination Inadequate

.

.

etc.

.

Application Study 1980 Diagnostic Checklist

Case Name _____
Your Name _____
Date _____

1_____Instant Word Recognition Adequate
2_____Instant Word Recognition Inadequate
3_____ Basic Sight Words Adequate
4_____ Basic Sight Words Inadequate
5_____ Sight Words Learned Via Decoding Adequate
6_____ Sight Words Learned Via Decoding Inadequate
7_____ Experiential Sight Words Adequate
8_____ Experiential Sight Words Inadequate
9_____ Visual Discrimination Adequate
10_____ Visual Discrimination Inadequate
11_____ Visual Memory Adequate
12_____ Visual Memory Inadequate
13_____ Print-Meaning Association Adequate
14_____ Print-Meaning Association Inadequate
15_____ Print-Sound Association Adequate
16_____ Print-Sound Association Inadequate
17_____ Other Adequate
18_____ Other Inadequate
19_____Decoded Word Recognition Adequate
20_____Decoded Word Recognition Inadequate
21_____ Sound-Symbol Association - Consonants Adequate
22_____ Sound-Symbol Association - Consonants Inadequate
23_____ Sound-Symbol Association - Blends/Diagraphs Adequate
24_____ Sound-Symbol Association - Blends/Diagraphs Inadequate
25_____ Sound-Symbol Association - Vowels/Vowel Patterns Adequate
26_____ Sound-Symbol Association - Vowels/Vowel Patterns Inadequate
27_____ Visual Segmentation into Syllables Adequate
28_____ Visual Segmentation into Syllables Inadequate
29_____ Auditory Segmentation into Syllables Adequate
30_____ Auditory Segmentation into Syllables Inadequate
31_____ Blending of Sounds Adequate
32_____ Blending of Sounds Inadequate
33_____ Adjustment of Blended Sounds to Language Adequate
34_____ Adjustment of Blended Sounds to Language Inadequate
35_____ Use of Root Word Adequate
36_____ Use of Root Word Inadequate
37_____ Use of Prefixes Adequate
38_____ Use of Prefixes Inadequate
39_____ Use of Suffixes Adequate
40_____ Use of Suffixes Inadequate
41_____ Auditory Memory Adequate
42_____ Auditory Memory Inadequate
43_____ Auditory Discrimination Adequate
44_____ Auditory Discrimination Inadequate
45_____ Visual Memory Adequate
46_____ Visual Memory Inadequate
47_____ Visual Discrimination Adequate
48_____ Visual Discrimination Inadequate
49_____ Other Adequate
50_____ Other Inadequate
51_____Meaning Vocabulary Adequate
52_____Meaning Vocabulary Inadequate
53_____ Number of words Adequate
54_____ Number of words Inadequate
55_____ Accuracy and Breadth of Definitions Adequate
56_____ Accuracy and Breadth of Definitions Inadequate
57_____ Other Adequate
58_____ Other Inadequate

Application Study 1980 Diagnostic Checklist

3

Case Name _____
Your Name _____
Date _____

59_____Oral Reading Adequate
60_____Oral Reading Inadequate
61_____ Instant Word Recognition Adequate
62_____ Instant Word Recognition Inadequate
63_____ Decoded Word Recognition Adequate
64_____ Decoded Word Recognition Inadequate
65_____ Rate Adequate
66_____ Rate Inadequate
67_____ Phrasing Adequate
68_____ Phrasing Inadequate
69_____ Intonation Adequate
70_____ Intonation Inadequate
71_____ Use of Syntax Adequate
72_____ Use of Syntax Inadequate
73_____ Use of Semantics Adequate
74_____ Use of Semantics Inadequate
75_____ Other Adequate
76_____ Other Inadequate
77_____Reading Comprehension Adequate
78_____Reading Comprehension Inadequate
79_____ Main idea Adequate
80_____ Main idea Inadequate
81_____ Paraphrasing Adequate
82_____ Paraphrasing Inadequate
83_____ Sequence Adequate
84_____ Sequence Inadequate
85_____ Cause-Effect Reasoning Adequate
86_____ Cause-Effect Reasoning Inadequate
87_____ Recall of Facts and Details Adequate
88_____ Recall of Facts and Details Inadequate
89_____ Other Adequate
90_____ Other Inadequate
91_____Listening Comprehension Adequate
92_____Listening Comprehension Inadequate
93_____ Main idea Adequate
94_____ Main idea Inadequate
95_____ Paraphrasing Adequate
96_____ Paraphrasing Inadequate
97_____ Sequence Adequate
98_____ Sequence Inadequate
99_____ Cause-Effect Reasoning Adequate
100_____ Cause-Effect Reasoning Inadequate
101_____ Recall of Facts and Details Adequate
102_____ Recall of Facts and Details Inadequate
103_____ Other Adequate
104_____ Other Inadequate
105_____Attention/Motivation Adequate
106_____Attention/Motivation Inadequate
107_____ Initiating Attention Adequate
108_____ Initiating Attention Inadequate
109_____ Maintaining Attention Adequate
110_____ Maintaining Attention Inadequate
111_____ Other Adequate
112_____ Other Inadequate

Application Study 1980 Diagnostic Checklist 4

Case Name _____
Your Name _____
Date _____

113_____Factors Effecting Learning Adequate
114_____Factors Effecting Learning Inadequate
115_____ Amount of Practice Adequate
116_____ Amount of Practice Inadequate
117_____ Attention Adequate
118_____ Attention Inadequate
119_____ Motivation Adequate
120_____ Motivation Inadequate
121_____ Visual Acuity Adequate
122_____ Visual Acuity Inadequate
123_____ Auditory Acuity Adequate
124_____ Auditory Acuity Inadequate
125_____ Transfer of Isolated Skills to Contextual Materials Adequate
126_____ Transfer of Isolated Skills to Contextual Materials Inadequate
127_____Other Adequate
128_____Other Inadequate

Appendix C

Explanation of Phi Correlation and Porter Statistic

Inter-Clinician Correlation

Phi

Given a domain for diagnoses/remediations/or cues (DX/RX or CX) for a given case, the Phi correlation is a measure of inter-clinician agreement. One Phi correlation is computed for each pair of clinicians.

The DX/RX/CX categories mentioned by one clinician are compared with those mentioned by a second clinician for the same case. This comparison is summarized in the table below.

CLINICIAN A, SIMCASE Y

CLINICIAN B, SIMCASE Y		PRESENT (+)	ABSENT (-)
	PRESENT (+)	Frequency count of statements in the domain present in both clinicians DX/RX or CX a	Frequency count of statements in the domain present in clinician B's session but not in clinician A's DX/RX/ or CX b
	ABSENT (-)	Frequency count of statements in the domain present in clinician A's session but not in B's DX/RX or CX c	Frequency count of statements in the domain absent in both clinicians' DX/RX or CX d

The calculation of the Phi correlation is derived from the table as follows:

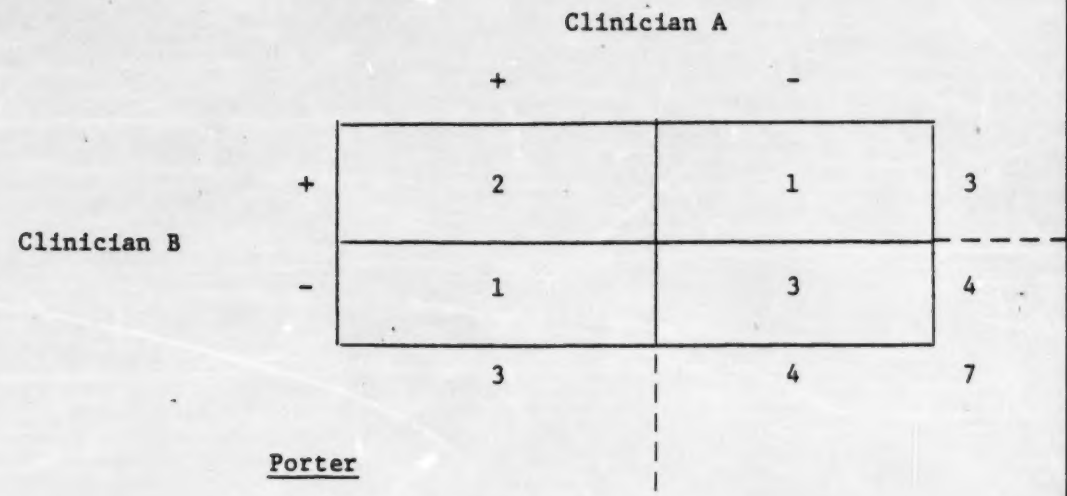
Clinician B, Simcase Y	Clinician A, Simcase Y: +	Clinician A, Simcase Y: -	
+	a(++)	b(+−)	a+b
-	c(−+)	d(−−)	c+d
	a+c	b+d	N

$$\text{Phi} = \frac{(axd - bxc)}{(a+c)x(b+d)x(c+d)x(a+b)}$$

The statistic is bounded by -1 (statements are in cells b and c only) and 1 (statements are in cells a and d only), only if the distribution in the marginals is equal. In all other cases the maximum and minimum values will be less than 1 and greater than -1.

An example of a completed table is as follows:

Statements of Clinician A, Simcase Y	Statements of Clinician B, Simcase Y	Domain of Statements
S1	S1	S1
S2	S2	S2
S3	S7	S3
		S4
		S5
		S6
		S7



Another statistic which is useful and based upon the 4-fold contingency table is the Porter Coefficient (P)* This coefficient is a measure of the agreement between two individuals' diagnoses or two diagnoses by the same individual as a function of the proportion of that agreement compared to the sum of the total agreements and non-agreements (i.e., the ++,+−,−+ boxes, and excluding the −− one.)

If A = no. categories common to both subjects or encounters,

B & C = no. categories mentioned by one subject (encounter) and not the other, then

$$P = \frac{A}{A + B + C}$$

This does not reflect agreement on not to select a category (diagnosis, cue, treatment) as is the case with the Phi correlation.

For the previous example, the Porter Coefficient would also be .67 as the total domain of categories is only 5. If the domain is large and the number of categories mentioned by both or either subject is small, the Phi and Porter values will differ, and the Porter coefficient will appear to be more reliable.

*developed by Dr. Andrew Porter, Institute for Research on Teaching, Michigan State University.